CA3 ONHW Q90
82 U6505 V.3

Pamphlet box

URBAN/MUNICIPAL

Review of Report of the Worker's health study of the Upper Ottawa Street Landfill study

Reference 36

CA3 ONHW Q90
82 U6505 V.3

Reference 36

APR 30 1987
GOVERNMENT DOCUMENTS

Review of

"Report of the Worker's Health Study of the Upper Ottawa Street Landfill Study"

by

S.D.WALTER, Ph.D.,

Department of Clinical Epidemiology and Biostatistics

McMaster University,

Hamilton, Ontario, Canada.

January 10, 1985

Introduction and terms of reference

This review is written in response to a request by Anne Koven, Research Director of the Upper Ottawa Street Landfill Site Study, in a letter to me dated December 13, 1984. That request indicates that the document to be reviewed is a draft. The written request does not give any specific guidelines for my review, and so I contacted Anne Koven by telephone for further details. In that conversation, I was instructed to carry out a general review of the document, with special reference to the methodology of data collection and analysis. Ms. Koven indicated that she wished my review to be critical in nature, and that it should emphasise weaknesses rather than strengths.

I think it is appropriate for me to indicate at this time that I have been somewhat involved in this project on previous occasions. This involvement has taken the form of a few consultative conversations with Dr Hertzman and Mr Mike Hayes, mainly concerning the most appropriate forms of analysis to be used with the data. I feel that this previous interaction with the Landfill Study team has been so irregular and peripheral, that it does not constitute a serious conflict of interest for me to now write this report.

Despite my mandate to be critical, I think I should state at the outset that the authors of this report deserve some congratulation on being able to carry out such their study with the thoroughness and rigour which quickly becomes apparent on reading the document. The political and social situation obviously made it difficult to carry out the perfect scientific investigation; yet despite the many external pressures on the team, they have demonstrated a high degree

of pragmatism, while attempting to maintain an attitude of impartiality and fairness to all parties. With the increasing public awareness of environmental issues, of the kind exemplified by the Ottawa Street Landfill problem, it is likely that epidemiologists will be called upon frequently to carry out many similar investigations in the future. While there are certain methodologic difficulties which have remained unresolved in this particular study, the general approach should serve as a useful role model for other health researchers faced with similar problems. With this in mind, I would urge the authors of this report to publish their findings as soon as possible, so that the valuable lessons learned here might be more widely appreciated.

Critique

The remarks which follow have been written approximately in the same order as they occured to me on reading through the text of the report. No other significance is to be attached to the order of my comments. Certain of my comments are marked with a "*"; these I regard as relatively major, in the sense that they may possibly affect the conclusions, or at the least the interpretation of some of the analyses. Some of these "major" problems might be resolved by further analysis of the data, but some are inherent problems with the design. The page number(s) for each comment refer approximately to a relevant section of the text of the report.

Detailed comments

1. (page 4).(*) Public concern about the landfill site was apparently raised in 1978. It is quite likely that there would be a difference, following this time, in the perception of health events as they occur, and also in the reporting of events at a later interview. It would have been useful to investigate this question, for example by comparing the data before and after 1978 with respect to over- and under reporting, and other validity measurements.

2. (page 4). Although I know nothing about the previous survey in 1979 carried out by the residents themselves, and although I would accept the authors' concerns about the lack of scientific rigour with which it may have been carried out, I nevertheless found it disappointing that there was not even a brief summary of the earlier findings.

3.(pages 12-13).(*). We are told nothing about how workers might or might not get onto employee lists. For example, are very short-term employees included, or is some minimum period of work required? This is important, because short-term people are less likely to be remembered by their co-workers.

Can we be assured that all longer term employees have been identified? The fact that some salvage pickers were identified by telephone contact with other salvage pickers, for example, suggests that the other methods of ascertainment were potentially incomplete.

Completeness of the employee lists is obviously crucial. We must feel confident that no-one has been missed, for instance, who dropped

out of this work group because of health-related reasons.

4.(pages 12-13).(*) It is a pity that the employers were not able to provide even a rough estimate of the amount of "exposure" of workers in the various job categories. Individual exposure estimates would be required by criterion 3 on page 7, but here (apparently) not even group estimates of exposure are available.

5. (page 14)(*). The selection of controls is also crucial to the study, as they are always the basis for comparison with the landfill workers. It appears as though the study team has had considerably less input into the selection of controls than they did for the exposed subjects. We have relatively little assurance, for example, that all appropriate control workers were in fact included in the sampling frame. Similarly we have to use faith in believing that the 50% sample of former outdoor workers was in fact chosen correctly. The denominator for this sample is unknown to the team (page 15). Finally, I have seen many examples of samples which were purported to be "random", but which on further investigation turned out to have been chosen in some rather non-random way. Generating a truly random sample is not difficult, but it does require care and an appropriate random number selection method. It would be highly desirable to have further details of exactly how these controls were selected.

6. (page 14) (*) The non-identification of deceased workers is an weakness. Unless all members of the working groups can be accounted for, the study ceases to be a complete study of the cohort. Excluding deceased workers from the control group quite possibly implies bias, especially if the reasons for these deaths were work-related.

7. (page 14). Table 2 is of very limited value, given the differences in age distribution which have been identified (page 13). Age-specific or standardised analyses would be much better.

8. (page 14). (*) The decision to only investigate the deceased exposed workers if the morbidity analysis was "positive" is questionable. This leaves open the possibility of an effect of a landfill effect on mortality, but not on morbidity, which would be completely missed. This strategy is especially troublesome given the rather high cumulative death rate (about 1/3) in the salvage pickers. This means that quite a high percentage of individuals are completely "missing" for what turned out to be the main analysis of the report.

9. (page 15). (*) Thinking of possible selection bias, what can the employers (or anyone else) say about the process by which employees are assigned to landfill versus other tasks. For example, do they have to be fitter, sturdier, or different in any other way for a landfill job, or conversely for one of the control job activities? Even if there are no conscious differences of this kind, are there any more subtle differences which might have crept in, and which might explain all or part of any health differences between the exposed and control workers?

10. (page 17). How far back did you go in the residential history, details of drinking water, etc? (Maybe one could tell by a close examination of the questionaire, but it should be stated in the text anyway).

11. (page 17). What attempts were made to validate item (j), i.e. the job history at the landfill?

12. (page 18) Even though cancer outcomes are rarer than some of the others, they are often of great concern in this type of investigation. It would be useful to have an idea of the expected number of cases of cancer of various types, even neglecting the 20 year latent period. Also can we then have a more precise idea of power, based on these expected numbers?

13. (page 19) (*) Why could the interviewers not be blinded, at least initially, with respect to exposure status? Could they not have collected the health information before going on to the job history (which would then unblind them, presumably)?

14. (page 20) (*) Why was it not feasible to do blind abstractions of all 120 health problems? Is this a fiscal problem, or one of some other kind? This is not clear in the text. Also how were the health problems which were analysed chosen from the 120?

15. (page 21) How is the percentage of consultation notes and discharge summaries which were found established? How do you know the denominator, consisting of all relevant consultations and discharges? Also, why not give the exact percentage "found"?

16. (page 23). Category 5 is tricky, and could almost be interpreted as non-confirmation or contradiction in the chest pain example given. If angina is ruled out, would that not mean that no chest pain could be confirmed?

17. (page 24) I suggest rewording "false" over-reporting to "apparent" over-reporting, here and elsewhere.

18. (page 24). (*). The main measure of validity is over-reporting. Can the team dispose of the possibility that exposed workers tended to report more symptoms and disorders at interview than the unexposed, but also did the same when seeing their doctors? If this scenario exists, one could have no over-reporting (and hence "validity"), but a strong difference between the exposed and the unexposed. This concern applies especially after about 1978, the beginning of public concern, but could also apply througout the time period if landfill workers tend (for whatever reason) to be more health-concious, leading to more reporting of symptoms at interview, more doctor visits, and so on.

19. (page 25). In what sense are the cumulative incidence rates "crude"?

20. (page 25) (*). What specific criteria were applied in step 1 of the analysis to "identify" important confounders? How bad did the difference between the health reporters and non-reporters (step 3) have to be before adjustment for confounding was made?

21. (page 25) (*). The criteria for a confounder, as given in reference #25, also include a requirement that the potential confounder be a predictor of the health outcome. (This is in addition to the condition quoted in step 3). The confounders actually included in the analysis rather dubiously satisfy both criteria; for instance in Table 13, analysis 5 of skin rash, adjustment is made for marital status. Is there any evidence that marital status is relevant to skin

rash?

Unless this second condition is met, overadjustment may be made for variables which are actually not confounders. The danger then is that such over-adjustment may reduce the apparent significance of the association between landfill "exposure" and health outcomes.

22. (page 25; step 4) (*) Concerning the person-years calculations:

(i)When did the counting of person-years begin for the non-exposed?
(ii)How was information on the exposed before their first exposure used, if at all?
(iii)What happens to workers who leave the landfill group - how are their person-years then counted?

Also, it would have been neat, had it been possible, to match the date and age of first exposure of the "exposed" to an appropriate control. This was not done because of the way the control sampling was carried out, but can there be any adjustment in the analysis for age and year effects of this kind?

Questions of this kind may have been dealt with already, but it is hard to say from the write-up.

23. (page 26). Only first incident events after "exposure" were counted here. Again, can there be any secular differences between the exposed and control workers?

24. (page 30).(*) Line g is a better summary of the completion rate than is line n; the latter could be very high if workers had a large number of doctors each, but only one chart was reviewed. Thus line n is potentially misleading in general.

In line m, "revised" should be "reviewed", I presume!

25. (page 31) Line 5: "Cases" should read "exposed workers". Ditto elsewhere on a number of occasions.

26. (page 31). Add sample sizes to table 6.

27. (page 31) (*). The t-test for the age difference between groups is inadequate as it only checks for a mean difference. Given the excess of young and old among the control group, the chi-square would be better, using appropriate age categories. Likewise, age stratification in some of the analyses might have been advisable.

28. (page 33 and elsewhere). (*) It would be a good idea to show the frequencies as well as the percentages in such tables.

Also tables with more than 2 rows (or columns) should be collapsed if small frequencies occur in some cells (e.g."other" as a source of drinking water).

Specific contrasts between rows (or columns) , or groups thereof, should be made to supplement the overall tests already done. At present the overall tests suffer from having possibly too many d.f., and hence reduced power.

Finally, why not quote the actual p-values in all cases, rather than just those which are significant? I do not like the footnote to this table - how much is "much greater"?

29. (page 32) Test for a difference in mean income, rather than using the multi-d.f. chi-square, again with possibly low power.

30 (page 38). Again, quote frequencies, and collapse as necessary (e.g. question 1d). Are the differences shown here significant (e.g.

question 1a; and question 4 - did the exposed start smoking significantly younger?)

31. (page 36) (*) The adjustment for exposure to toxic substances not at the landfill is described as being "conservative". This is not true, as seen in the following two examples.

Example 1

	Exposed to toxins elsewhere		Not exposed elsewhere	
	Symptom	No symptom	Symptom	No symptom
Landfill workers	200	50	500	250
Control workers	560	190	150	100

In the workers exposed who are exposed to toxins elsewhere, the odds ratio is 1.36; in those not exposed elsewhere the odds ratio is 1.33. Hence an analysis which adjusted for exposures elsewhere would indicate about a 33% increase in the odds of a symptom being present among landfill workers, relative to controls.

If the other exposures are not controlled in the analysis, the aggregated 2x2 table becomes

	Symptom	No symptom
Landfill workers	700	300
Controls	710	290

with an associated odds ratio of 0.95, or about 1. Thus the controlled analysis shows a greater effect of landfill exposure than does the uncontrolled analysis.

In the second example, the reverse occurs.

Example 2

	Exposed to toxins elsewhere		Not exposed elsewhere	
	Symptom	No symptom	Symptom	No symptom
Landfill workers	80	320	400	200
Control workers	68	532	200	200

In the workers exposed elsewhere, the odds ratio is 1.96; in those not exposed elsewhere the odds ratio is 2.00. Hence an analysis which adjusted for exposures elsewhere would indicate about a doubling in the odds of a symptom being present among landfill workers, relative to controls.

If the other exposures are not controlled in the analysis, the aggregated 2x2 table becomes

	Symptom	No symptom
Landfill workers	480	520
Controls	268	732

with an associated odds ratio of 2.5. Thus the controlled analysis shows a smaller effect of landfill exposure than does the uncontrolled analysis.

In my opinion, a situation like Example 1 is the more likely of the two. In that example, the proportion of all workers reporting symptoms is greater among workers who are exposed elsewhere, which seems plausible if such exposures are indeed related to health. If a situation like example 1 applies in practice, then control of such

extraneous exposures is liberal, i.e. it will tend to attach too much significance to the health effects of landfill work.

I find it very difficult to interpret the analyses after adjustment for these other exposures. It would have been better, if possible, to control for such extraneous effects by an appropriate (e.g. matched) sampling design. As it is, the landfill workers obviously have less chance to report such other exposures (as is true, incidentally, in my example 1), and this appears to adversely affect the analysis of the landfill effect.

32. (page 39) Is there any explanation of why the percentages exposed to metal fumes are so high, especially in the landfill workers? This seems to imply, at least, that almost all workers in both groups had had some other job at some time.

33. (page 40). (*) A repeat query, but why did you not analyse all 120 conditions? One could actually have the situation where the cumulative (crude) incidence is equal in the two groups, but a significant difference becomes apparent after adjustment for confounders. Confounding can either enhance or mask a landfill effect, and your strategy appears to run the risk of missing situations of the latter kind.

34. (table 13) Some of the p-values will be rather approximate because of small sample sizes.

35. (page 45) . The analyses described in the third paragraph could be further improved by using the Cox model with time-dependent covariables. For example, one natural covariable in the context would

be the amount of dumping carried out at the landfill in the previous year - this would be a good way to investigate short term effects. Another possibility might be the cumulative exposure of a given worker at any given point in time - this would be a good way to evaluate the dose-response pattern, if any.

36. (page 45, line 3 from end). "risk" should be "relative risk".

37. (table 14). You could quote actual rates in this table, which would also avoid the problem of zeros.

38. (table 14). I am still not clear exactly how the person-years have been counted for this table; do the calculations include person-years only during each time period, or all p-y up to that time?

39. (Table 14). It might also have been helpful to examine each condition in each time period in the way shown, rather than just giving one time period for each. I presume what is actually shown is the period of highest relative risk; if so, why not show the other time period rates for comparison?

40. (table 14). Are the relative risks shown significant?

41. (page 47). Some might say that statement 2, and similarly elsewhere, prejudge the issue. Why could we not, for instance, envisage the possibility of delayed "tremor" effects, which seem to be "ruled out" as a reasonable possibility at this point in the text.

42. (table 15). Again I am not sure which person-years are included here.

43. (page 49). The first sentence of statement 2 is obscure to me.

44. (page 49). I agree with point 3, especially as we apparently do not know the reasons for leaving employment.

45. (page 50). Are the results in table 16 age adjusted, or adjusted for any confounders? If so, say so!

46. (page 54, and elsewhere). "Clustering" seems like the wrong term here. In line 10 from the bottom, for example, I would suggest replacing "clustering" by "greater dependence".

47. (page 55). (*) To answer the question "have health status differences which emerged after first exposure to the landfill, persisted to the present", it seems to me that one should examine changes within the same individuals, rather than the "unlinked" group analyses of Table 18, for instance.

48. (page 55) (*) The analyses of "mood" symptoms would, I suspect, be especially suceptible to the effects of "hysteria bias" induced by the general public concern recently.

49. (page 58) (*). To say that there were "only" two birth defects among exposed workers is rather casual; could we at least know the expected number of defects, based on the number of pregnancies, ages of mothers, etc.

50. (table 20) (*) It is well known that proportional mortality analyses are relatively weak. As you already know the person-years at risk, why not compute the SMR's, based on the Ontario rates? You could do this in addition to the present PMR analysis; the latter is useful in avoiding the effects of the " healthy worker effect", which would

clearly affect the comparison between this group of workers and the entire provincial population.

51. (table 25) Again some collapsing of some small frequencies (here columns) seems in order. For example, if the "not confirmed" respiratory responses are combined with the "possible confirmed", you would have a test with only 1 d.f. instead of 2 d.f. In this example, the contrast is then significant at the usual 5% level. I suspect that many similar examples like this could be found in various tables of the report.

52. (page 64). (*) The second full paragraph seems to contradict the data in Table 24, where a trend is evident.Has an efficient test been applied specifically for trend, or has the usual multi-d.f. chi-square been used again? It would be interesting to carry out this trend analysis by type of condition or symptom, especially for conditions/symptoms which might be more easily forgotten (for example less "serious" conditions).

53. (table 26). (*) The footnote indicates that Fischer's (sic) exact test was used, apparently on 2 x 3 tables. Although exact tests exist for tables larger than 2 x 2 (on which Fisher's test is appropriate), one must be careful to specify which tables are more "extreme" than the one observed, and there are a number of possibilities here. What exactly has been done in this table?

54. (page 75). (*) The argument at the end of the first paragraph is not quite right, because the chance of a type I error is always fixed in advance (e.g. at 5%), regardless of the sample size. Therefore it

is wrong to "eliminate" the 10 comparisons of low power, because they too had that 5% chance of giving a "false positive" result.

The issue of multiple comparisons is very tricky because the usual adjustment of the p-value may be unduly restrictive, IF the study is indeed to be regarded as hypothesis generating. If however the data may be used for policy decisions concerning the workers now, i.e. before the residents' study has a chance to confirm or deny these findings, then you should be strict on the adjustment, and effectively require that an association in this study be very strong (and significant) before acting.

55. (page 75) In the second paragraph, your arguments about the nature of the adjustment for confounders require, as I indicated earlier, a demonstration that they are indeed predictive of the health outcome in question. Thus for example, you should provide evidence that marital status is related to the frequency of mood disorders, and that exposure to fumes and gases is related to skin disorders. And in order to be able to say in which direction the adjustment would go, you need to know the direction of the association of the confounder with the health outcome.

56. (page 78). Where have the conditions listed here come from? Why were the adult conditions shown here not dealt with more fully in the workers' study?

57. (general). You may have said, but I did not recall a precise description of when the interviews were done. Which year(s)? And which season(s) - I would imagine that symptom reporting could be influenced by the month of interview, and that this could even affect the

reported data for earlier years.

58. (general) As noted earlier, you might be more specific in the text as to the exact period of recall required for study subjects, concerning health history, and history of residence and other factors.

59. Last but not least, the title of your report is very cumbersome!

Good luck with the residents' study!

Stephen J. Walter

S.D.WALTER

January 10, 1985, Hamilton.

CCT Canadian Centre for Toxicology
Centre Canadien de Toxicologie

January 15, 1985

Dr. Anne Koven
Research Director
Upper Ottawa Street Landfill
Site Study
20 Jackson Street West
Suite 412
HAMILTON, Ontario
L8P 1L2

Dear Dr. Koven:

I appreciated the opportunity to review the workers' health study of the Upper Ottawa Street Landfill Site. Overall my impression is that the authors have paid careful attention to detail and have done as good a job as is possible given the difficulties associated with this sort of project. I have several points I would like to make that I think might strengthen the report.

1. It is important that scientific peers reviewing this report be given access to as much information as possible about conditions at the dumpsite. Hence I think it is important to describe conditions there with regard to work functions performed and opportunities for exposure. Some further description of the variety of tasks performed would be beneficial. In addition, and probably more importantly, I believe it is important to include in the document, possibly as an appendix, a listing of the major classes of chemicals found in the dumpsite to date. This data, coupled with the description of the potential range of toxic effects ascribed to these substances will lend credability to the authors conclusions.

2. While the association between landfill site exposure and health problems specified in Table 28 appears to be positive one is left with a nagging doubt about the strength of this association. While I agree the authors have taken extraordinary precautions to ensure study biases were considered, inadequate attention is given to providing the reader with details concerning the workers' symptoms. For example, did the workers' visit their physician more often? Where there health complaints by the workers? Did the workers express views regarding potential health effects of dumpsite chemicals? Did the authors distinguish clearly between health complaints/ symptoms and physician diagnosis of health effects in Analysis 3?

.../2

3. Any study of this nature is bound to raise great public/political alarm. It is therefore critical that the authors spell out in every detail the limitations of the study design and conduct and the impact these have on the conclusions. I believe these qualifying statements need to be included in the summary and conclusions sections.

4. While the initial intent of the study was hypothesis generation there can be no doubt that the authors will be required to draw conclusions about the health status of the workers. It is therefore my contention that the conclusions and summary should include such as discussion, subject of course, to the caveats expressed in Item 3 above.

5. Finally, some minor points in the text.

   a) p.11. The term subclinical is misused here. Surely mutagenesis, neurotoxicity, hepatotoxicity and metal toxicity are not subclinical effects. I would suggest the term "measures of" rather than subclinical.

   b) The Tables use various descriptors to identify health effects observed (e.g., narcotic symptoms = headaches, dizziness, etc.) These should be standardized throughout the Tables and text.

   c) p.55. Table 19 should read Table 18.

May I take this opportunity to wish you the best in completing the report. A summary of costs associated with this review is attached.

Yours sincerely,

I.C. Munro

Ian C. Munro, Ph.D.,
Director

ICM/ec
att:

Review of the REPORT OF THE WORKERS' HEALTH STUDY OF THE UPPER OTTAWA STREET LANDFILL SITE STUDY

Philip J. Landrigan, M.D., M.Sc., D.I.H.
National Institute for Occupational Safety and Health
Cincinnati, Ohio, U.S.A.

This was a questionnaire study intended to assess the occurrence of adverse health effects in 176 workers who had occupational exposure to the Upper Ottawa Street Landfill site in Hamilton, Ontario. The study also included, as a reference group, 235 municipal workers who had no occupational exposure to the landfill. The Upper Ottawa Street Landfill is a large dump-site which received all varieties of domestic and industrial wastes from the early 1950's until it was closed in 1980; more than 100 chemicals have been tentatively identified in the site.

The study population consisted of four sub-groups--two groups of landfill site employees, salvage pickers, and firefighters. The reference population consisted of road crews and cemetary workers. The groups were not individually matched. All participants in both groups were administered a long pre-tested questionnaire in face-to-face home interviews by trained interviewers. The questionnaire sought information on demographics, occupational exposures, and a wide variety of health conditions. It also included "decoy questions" to distract participants from items related to environmental exposures, and it included certain questions on chronic disease from the Canada Health Survey to serve as a "benchmark" against

national data. The study protocol also made detailed provisions for the review of family physicians' medical records for all participants who reported adverse health conditions. The main analysis of the data was actuarial in nature and was performed using the Cox proportional hazards model.

Response rates to the interviews were very good among exposed workers, 84.5%, and acceptable in the reference population 71.9%. Also, 86.3% and 86.8%, respectively, of interviewees in the two groups consented to medical record abstraction. The groups were found to be quite comparable in age, education, income and drinking habits. Some differences were evident in sociodemographic variables, including marital status and exposure to well water, and in cigarette consumption.

The analysis of interview data was carried out three times. Analysis 1 was the most sensitive and least specific. It included all positive questionnaire responses without regard to whether they had been verified by chart review. Analysis 3, on the other hand, was the most restrictive and included only medically verified conditions. Strong associations were found between landfill exposure and both chronic bronchitis and mood alteration. A moderate association was found between landfill exposure and skin rash. Efforts to assess date of first occurrence of these conditions were not altogether convincing, but suggested that the highest incidence for rashes was from 1965 to 1969, and for bronchitis and mood alteration from 1970 to

1975. The latency interval from first exposure to appearance of each of these symptoms was reported to be 0-5 years. No particular correlations amongst symptoms were noted; there appeared to be no identifiable "landfill syndrome".

A proportionate mortality study was conducted. It was, however, based on only 21 deaths, and results were inconclusive.

Comparison of self-reported health data with results from the Canada Health Survey showed that, in general, reported prevalences of most symptoms were higher in both the exposed workers as well as in the comparison groupthan in the general population of Canada. These findings appear to reflect differences in interview structure and in the formatting of questions and also possibly differences in the level of sensitivity. These findings emphasize the need for concurrent controls in epidemiologic studies of emotionally charged issues.

To assess the quality of self-reported interview data, interview results were compared with the results of medical record verification. Overall, 88.9% of reported illnesses in persons who reported having seen a doctor were either definitely or possibly confirmed by medical records.

My overall impression of this study is that it was extremely carefully planned and meticulously executed. The authors were aware that this study would be viewed as a prototype for other landfill studies in Canada, and they sought to provide highly credible results. In my opinion they have succeeded.

The several features of this study which give it such great credibility are:

- the very careful design of the questionnaire;

- the provision for systematic followup through medical record review of all reported positive results;

- the verification of results through comparison of data obtained from various sources; and

- the generally careful and understated manner in which [redacted] results are presented.

Also the overall approach of the study, which began by evaluating the most heavily exposed segment of the population--namely, the workers exposed occupationally to the landfill--makes good sense. It seems eminently reasonable to me that examination of heavily exposed workers should precede and guide the subsequent assessment of less heavily exposed community residents.

I congratulate the investigators on a job well done.

REVIEW OF THE UPPER OTTAWA STREET LANDFILL SITE STUDY

BACKGROUND

In 1978 (according to the report under review) public awareness of possible adverse health effects resulting from exposure to dumpsite contents was spurred by the Love Canal Incident. An informal survey carried out by residents in 1979 suggested the existence of health problems. In 1980, a Committee was formed to investigate the chemical nature of dumpsite materials and to explore the possibility of health effects engendered by such exposures. The authors of this report have correctly acknowledged that epidemiologic studies benefit from knowing the nature, quantity and routes of exposure(s), as such information facilitates the selection of valid, sensitive and reliable measures of health status. Unfortunately, few data were available, which is usually the case for dumpsites. The investigators apparently expected some criticism about undertaking a study without detailed environmental data, as they justified the importance of such studies at length.

DESIGN

It is most laudable that the investigators adopted a strategy of empirical hypotheses-generation as Phase One of the studies to be undertaken. A retrospective cohort study was used, which is entirely appropriate given the history of disposal and the closing of the dump in 1980. The decision to undertake the study among workers who were presumed to have higher exposures than residents is also reasonable. However, as no data on duration of employment was presented and as no data on exposure is available, long-term residents may have a greater total exposure than workers.

DATA COLLECTION

In studies such as these, data may be collected through: i) self-report from study subjects; ii) records review; and iii) clinical examination. The first two methods were used in this study, with the records review serving as a means of questionnaire validation rather than an independent data source.

QUESTIONNAIRE CONTENT

The questionnaire development seemed comprehensive as described in page 11-12 and 15-17. The questionnaire contents were correctly described in the report. The investigators appeared to have taken into account questions used in similar investigations, residents' concerns as well as plausible health effects engendered by a list of 100 chemicals. Since this list was not provided in the report, one cannot verify the completeness of the health effects included in the questionnaire. Furthermore it is likely that these effects are only suggested by studies of animals and high occupational exposures. The non-specificity of the outcome measures increases the chance of alpha errors, a possibility which the investigators thoughtfully acknowledge. Given that percutaneous and inhalation exposures were

anticipated, measures of dermal and respiratory toxicity were appropriate. Since many industrial chemicals are neurotoxic, neurobehavioural and frank neurotoxic outcomes were also appropriate. Given the seriousness of cancer and adverse reproductive outcomes, these are also pertinent topics, although the present sample size precludes adequate statistical power to draw definitive conclusions. Pretesting was minimal, but the questionnaire appears to have been professionally constructed. Note that there were no questions about duration (apart from what could be inferred from the beginning date of an effect and its continued presence at the time of interview) or severity of the conditions of interest. The lack of subclinical outcomes is justified by the relative underdevelopment of study hypotheses as well as by a lack of knowledge about the validity (i.e., sensitivity and specificity) of such measures. Interview conditions appear to have been comparable for both exposed and unexposed groups. Unfortunately, interviewers were not blind with respect to exposure status. It is not clear why the exposure status could not have been concealed until the end of the interview by rearranging the sequence of certain items.

MEDICAL CHART ABSTRACTION

This exercise was undertaken to identify possible recall bias -- i.e., as a method of collected actual clinical data. Based on the data presented on page 30, records were abstracted for only 70% of the the total eligible among the exposed, or for 82.6% of those interviewed, and for only 52% of the total eligible among the non-exposed, or for 72.6% of those interviewed -- making the results very vulnerable to bias. Although nurse abstractors were blind to exposure status, they were looking specifically for problems mentioned in the interview, an approach which is more likely to detect overreporting than underreporting. How far back in time any record extended (usually the record of one physician only per subject was abstracted) was not clear. Neither was it explicitly stated whether nurses verified that conditions complained of during a certain time period occurred within the time period noted by the physician in the medical record. On page 24, the investigators claim that, "For the purposes of this study, it has been assumed that false overreporting was equally distributed between exposed workers and controls; that the quality of record keeping was equally high among the physicians of each group." In fact, the investigators did not make the first assumption, because it was subjected to analysis in Tables 25-26. The classification categories for complaint confirmation developed for nurse abstractors are an example of overkill. One has to wonder whether the time and effort spent in this endeavour might have been better spent on the collection of some actual clinical data. This scheme may be further criticized on the grounds of validity: of the four "possible" categories only #4 seems to represent a truly possible documentation.

It is not clear from the report how the physicians were identified whose records were reviewed for purposes of validation. If they were identified primarily on the basis of reports by the respondents themselves, then there is a high probability of recall bias with respect to the identification of attending physicians themselves. If physicians' data is to be used for validating purposes there is no shortcut, in my opinion, to the use of

provincial health insurance records to search the names of all possible attending doctors in defined periods of time. Why death certificate information alone was used for deceased workers is unclear. Further, could not clinical records, particularly those pertaining to the last episode pertaining to death, have been obtained? They could have been quite revealing.

In page 20 the statement "self-reported health problems were confirmed by abstracting the family physician records of all study subjects who signed a consent form" (my emphasis) needs clarifying. Does the family phyician mean one main attending physician or is this a generic term?

STUDY AND COMPARISON COHORTS

The description of the study and comparison cohorts is insufficient to conclude that these samples were either exhaustive or representative of the underlying populations. It is likely that deceased and short-term workers may not have been identified as readily as long-term workers who were living at the time of the study. While the choice of an external comparison cohort of other outdoor workers was reasonable, it is unclear why a sample of retired non-exposed workers was included. This inclusion resulted in non-exposed workers being younger and older than exposed workers, although this is obscured by the presentation of mean ages only. Thus, bias may have resulted from: i) the response rates discussed in the previous section; ii) the possibility that short-term and deceased workers were less likely to have been identified; iii) the difference in age structure between the two groups (which would have been associated with duration of employment, a variable which is conspicuously absent from this report). Although the investigators attempted to demonstrate that the two samples were comparable on a variety of sociodemographic variables, there was no analogous comparison between respondents and non-respondents.

One must assume that all workers were white males, although this is not explicitly stated. In this situation one would have preferred an analysis matched on age (which is related to most health outcomes), duration of employment (as a proxy for exposure), time of beginning employment (since exposure as denoted by dumping contents and quantity was inconstant over time), and current employment status.

The authors stated that a matched analysis was impossible since the ages of the respondents were not known until the survey was done. Nevertheless, some information could have been obtained from all eligibles which would have permitted post-survey matching or stratification on employment duration as well as an analysis of the comparability between respondents and non-respondents.

Finally, some of the study cohort and an unknown number of the non-exposed continued to be exposed after 1980 to their respective conditions, thus invalidating any post-1980 comparisons.

ANALYSIS

Comparability of Sociodemographic Variables

Although the exposed group was more likely to be foreign born, better educated, married, lighter drinkers and heavier smokers than the non-exposed group, these differences were not significant.

Main Analysis

The Cox proportional hazards model replaces logistic regression when using an incidence density instead of a cumulative incidence approach. In this case, person-years are only counted until the first event occurs. This means that the earlier one gets the condition the smaller the denominator will be in comparison to the numerator. Thus, the relative risk will depend not only on the occurrence of the events, but on the time of occurrence. The analysis of first occurrence only, as was done here, is fair only if the two groups are comparable with respect to salient variables such as age, duration of employment and time of starting employment.

Three analytic methods were employed. From the description given on page 26, it seems as if events were included in a given analysis based on their validation through medical charts. No mention is made of how events were analyzed for the 20% or so of respondents whose charts were not abstracted, and one cannot guess as no total sample size is given in Table 13 (the main analysis). Furthermore, the difference between analytic methods #1 and #2 is unclear, although one might guess that method #1 accepted all events except those in category #8, while method #2 accepted all those except those in categories #7 and #8. Attention should be focussed on method #3, the least powerful but the most unbiased method. Table 13 shows elevated relative risks among the exposed for respiratory symptoms, skin symptoms, mood and "narcotic" symptoms. The rationale for combining certain symptom clusters was not established a priori, which leads one to suspect that certain conditions might have been combined until significance was reached. The only skin conditions which met the criteria for reporting was rash ($p=.083$, 1-sided). However, by combining rash, unusual acne (probably included to embrace chloracne, which results from certain industrial exposures and would be unlikely in this study) and discoloured patches on skin (liver spots?) significance was reached.

It would have been illuminating to analyze the period prevalences for the groups compared in two ways: i) by restricting the numerators to symptom or health problem events elicited spontaneously before specific probes for selected entities was made (i.e., questions #9 and #15 in the questionnaire); ii) by allowing all events to be counted in the numerator whether they had been elicited spontaneously or by probe.

No association was found between exposure and certain decoy conditions (page 44). However, one might have expected some gastrointestinal or haematologic disturbances, which were not found.

Descriptive Analyses

Although not explicitly stated, the relative risks presented in Tables 14 and 15 appear to have been calculated from a 2x2 table. These risks are not adjusted, and one cannot tell whether all self-reported events, or just confirmed ones, were included. Table 14 might have benefited from the inclusion of all time periods for each condition listed. This table is not accompanied by a critical interpretation. For example, why would the highest relative risk for mood occur in 1970-1975, while that for "narcotic" symptoms occur in 1976-1980? And why would the relative risk for red eyes, which one would expect to be an acute condition, occur after the exposure had ceased?

With respect to Table 16, the p-values should probably be higher than what is reported. The correct analysis would be to partition an overall chi-square with 2 degrees of freedom into an exposed vs. non-exposed comparison and a between exposure groups comparison. Because this seems to be an a posteriori comparison, one should probably compare the test statistic to a chi-square distribution with two degrees of freedom. There are few differences between the two exposure groups which does not support a dose-response relationship. Inexplicably, one of the most significant differences between the two groups concerns muscle weakness, which is not a common effect.

The conclusions regarding Table 17 are justified: there does not seem to be a "landfill syndrome".

Table 18, which concerns concurrent prevalence, seems to be based on inference tests using logistic regression. Again there is very little critical interpretation of these data. For example, difficulty breathing is no longer significant although chronic bronchitis is. Could this difficulty in breathing have been due to asthma? cardiac problems? emotional distress?

The investigators presented different data to demonstrate that overreporting among the exposed did not account for the results. Table 19 suggests that the exposed were more likely to have received Workman's Compensation than were controls even though both groups reported the same frequency of work-related injury. This may suggest that the exposed suffered more severe injuries. Alternatively, this might suggest that the exposed more readily perceived and complained about such injuries. The comparison with a blue-collar sample on part of the chronic health problems segment on the Canada Health Survey revealed that both exposed and non-exposed were less healthy than their counterparts nationwide. There is no mention of age adjustment for these data. The fact that the exposed had significantly more hearing disorders, sight disorders and dental problems (which might not be expected to be associated with dumpsite exposure) than did the national sample while the non-exposed did not, might suggest a tendency among the exposed toward poor health in general or toward overreporting. Moreover, there were few differences between the study and comparison cohort on this segment of the questionnaire (Table 22).

The results of the main analysis might well have been produced by underreporting among the controls as well as overreporting among the

exposed. Unfortunately, Table 23 presents overall confirmation rates without breaking them down by exposure category, so that the possibility that underreporting among the non-exposed influenced the results cannot be examined.

In Tables 25 and 26, chi-square analyses were performed to determine whether there were differences in the proportions of "confirmed", "possibly confirmed", and "not confirmed" by exposure category. It appears (because Fischer exact tests were done for some conditions) that two of the categories were probably collapsed, which would be incorrect given the data presentation in the table. It is not possible to discern exactly what was done given the poor description of statistical methods already noted. However, it appears that some overreporting among the exposed was likely for respiratory problems and "narcotic" symptoms.

Table 27 argues against exposed subjects being hypochondriacal.

CONCLUSIONS

The project is adequate for purposes of hypothesis generation. The clusters of clinical health problems identified for hypothesis testing seem justified and are plausible. No firm conclusions about the landfill site as a "risk factor" can or should be drawn even for the workers. That is because the methods employed in this study leave room for selection bias (sample composition and participation rates) and misclassification bias (interviewer knowledge of exposure status, possible overreporting, probing questions) all of which have been discussed at length above. Analyses of disability data (taken from the Canada Health Survey instrument) and participants' opinions regarding their perception of the role of exposure in their self-reported conditions are conspicuously absent. From a perusal of Table 13 one must be concerned about skin disorders, respiratory disorders, mood disorders and "narcotic" symptoms. As far as skin disorders go, the main problem is rash, which is not persistent following cessation of exposure. The evidence for respiratory problems is conflicing: there are no differences between the two groups on the Canada Health Survey questions (which precede later probes) with respect to bronchitis and emphysema, although there are for "other" respiratory problems (whatever that is interpreted to mean). This is in contrast to the main data analysis of specific probes. It should also be noted that the tendency toward overreporting of respiratory symptoms by exposed subjects was of borderline significance ($p=.06$). Table 26 provides evidence of overreporting of "narcotic" symptoms among the exposed. Yet the exposed sample did not differ significantly from the Canada Health Survey sample in the reporting of mental disorders, while the controls were significantly less likely than the national sample to report mental disorders. Thus, the mood and "narcotic" symptoms reporting may reflect the better than average mental health of the control sample. It is noteworthy that although neurobehavioural symptoms (which may have causes apart from nervous system pathology) gave cause for concern, there was no evidence of frank neurotoxicity as demonstrated by neuromuscular problems.

The investigators' conclusions that these results will be treated as probable high dose effects (page 79) require further consideration. I feel very strongly that to proceed in that manner is not only unethical but likely to be harmful to the population. One cannot declare that one is engaged in a hypothesis generation study and then reach conclusions that contradict the stated a priori purposes. Although the methods employed and the reasonable compromises made in the study are adequate for the generation of hypotheses, they are grossly inadequate to reach any conclusions that would be invoked in health policy, or public policy formulation, or implementation of policy. I should like these remarks on my concerns about the ethics of using these results for policy to be a matter of pulic record. One cannot object to the use of the four identifed problem areas for use in the residents' study, although one might hope that the instruments could be further refined; e.g., the incorporation of a questionnaire validated for epidemiologic purposes such as that used by the American Thoracic Society. One might also hope that clinical and subclinical components could be incorported into this study; e.g., simple spirometry, examination for skin conditions, and neurobehavioural tests to assess such parameters as reaction time, memory and mood. Further recommendations for a future study with residents include the need for much more stringent rigour and blinding data-gatherers and data-handlers in the entire process of field work and analysis. It includes much more exhaustive methods of following of subjects to obtain high and very comparable response rates in all comparison groups employed. I found it surprising, given the existence of many well-known and well-used questionnaires at least a decade old in the McMaster/Hamilton research community, that their format was not used so that comparisons could be made with unselected probability samples of general populations from the same area. At the moment there are also questionnaires available from the Junction Triangle Study in Toronto (in four languages), large portions of which could serve purposes of a resident study in the Upper Ottawa Street area in the future thus affording comparisons with unselected populations of the same province. A most important recommendation is the judicious use of spontaneously elicited symptoms and health problems as well as those obtained through probes. In earlier work done by this reviewer in Smithville we demonstrated empirically that you could increase the frequency of clusters of symptoms by a factor of seven just by changing from spontaneous data-gathering to probe data-gathering. The sensitivity of health problems and symptoms to the tactics of asking questions is so high that the order of magnitude of differences would drown those shown in this preliminary study now under review.

Opinion of This Hypothesis Generation Study

This study is an empirical effort much better than average for generating hypotheses to be tested later in definitive study. I reiterate most strongly that the use of this information should be restricted stringently to the purposes for which it was designed, hypothesis generation. Invoking the information for policy or for conclusions about the workers or residents is not only invalid but it is unethical because it could be very harmful. There are many biases up to this point that cannot be ruled out. The biases

assume special importance when the order of magnitude of differences between the groups compared is not particularly large.

Walter O. Spitzer, M.D.
McGill University
January 1985

TYPED ON FEB. 14 BY TERRY WALKER

REVIEW FROM MARVIN SCHNEIDERMAN

February 4, 1985

Dear Miss Koven:

Enclosed is my report, as requested, concerning your draft report of the workers' health study. I'm sending you a hand-wirtten copy to reduce the delay that might be occasioned by having the material typed. I think you will find my handwriting legible.

My approach was as follows: I did a first reading of the report and marked it in detail shortly after I had received it - about December 21 or so. Some issues concerned me, and I wanted to think about them - so I put the report aside. I then re-read the report late in January, re-examined the issues that concerned me, did a modest number of computations (some of these appear in my review) and then prepared a written report.

The written report follows this form:

1. General review
   - strength
   - weaknesses
   - some issues of concern

2. Comments on details
   - alternative analyses and related suggestions
   - questions on problems in the text - issues of understanding

I spent two full working days on the review and the report. How shall I bill you for this?

I hope you find my analysis helpful.

MARVIN SCHNEIDERMAN

SUBJECT: Report of the Workers' Health Study of the Upper Ottawa Street Landfill Site Study (undated, late 1984) Hertzman et. al.

1. GENERAL

This is a careful and thoughtful description of a difficult epidemiologic problem. The authors have been concerned with many issues involved in conducting such a study, namely; 1) appropriate choice of controls, 2) bias-particularly recall bias, 3) measures of health (or ill health) and their verifiability, 4) exposure - both aspects of intensity and of duration, 5) biological reasonableness of the findings, 6) confounding factors (or possible confounding factors), 7) power to detect effects (at times), 8) informants for interviews, with emphasis on first-person informants.

In general one would say that serious attention had been paid to these problems and, in most cases they were dealt with as well as could be in view of the difficulties such a problem presents.

The authors summarize their study in the light of its potential for providing hypotheses to test in future studies. This is both a strength and a weakness. In the early portions of the text the impression is given (to me at least) that the main concerns were to have been addressed to actual health effects upon the workers, and that hypothesis testing (or really hypothesis generating) was of lesser concern. Re-writing of the introduction, in which hypothesis generating appears now as the third item - after "current and past health problems", and "excesses in morbidity and mortality" - and placing hypothesis generating first, would make the beginning and the end of the report more consistent with each other.

I am not sure, however, that the substantial conservatism implied by this emphasis on hypothesis generating is wholly warranted. While the authors are to be commended for their modesty, and their attempts to be rigorously scientific, they may have been too modest

or too conservative. Late in the report a table appears (Table 28) which strongly suggests that more than hypothesis generation was possible from the data collected here. The column "biological plausibility" implies to me that a body of prior knowledge is likely to have existed before this study was undertaken that suggested that some effects (perhaps 11 out of the 13) were, or could have been anticipated - either from report of putative effects seen in persons exposed at other dump sites, or from knowledge of the materials, or some of the material contained in the Ottawa Street site and their likely toxic effects. If this was the case, then it is unfortunate that the anticipation of these health effects was not set forth before any further data were collected - so that there might not, or should not, arise the issue of data dredging, or of multiple testing. This proposed mode of action is now set forth in the conclusions (page 78, 79), involving severe items that a workers' study could not possibly have led to, or generated hypotheses for. These include most of the childhood manifestations to be looked for, and the female reproductive problems.

Nonetheless, this study has substantial strengths, and it is somewhat regretful that not setting out specifically what might be expected to be found, may have reduced the true worth of the study.

The specific strengths that I see are these:

1. The use of first person interviews. The informants were the actual persons at risk. If there are important "subjective" effects, gathering data from persons other than those at risk could lead to distortions.

2. The use of medical records to verify or confirm personal reporting. The results of the nurse-analysts might have been strengthened by having independent second reviews of a random set of the records - but this is not a serious failing. (Quality control of this type is expensive - and funds might be better spent some other way).

3. The seletion of controls (other "outdoor" workers) was good - although some special break-downs in the analysis might be in order in view of the bi-modal nature of the age distribution of the controls.

4. The attempts at developing (relative) exposure estimates are worthwhile. (In this regard, I would have liked to have seen more of the results reported in terms of relative exposures. Table 28 has a column headed "Risk Gradient by intensity of exposure", but I saw little of this in the body of the report. One might have reported by duration x intensity as a possible exposure measure).

5. Comparisons with the Canadian Health Survey were worth doing - if for no other reasons than that the outcome showed how difficult it is to do truly comparable surveys, thus stressing the need for an internal control.

6. The discussions of potential biases, confounding, etc. are generally good - although I could not follow one late discussion of power (bottom of page 73 and top of page 75) I could not see where the numbers came from, or what they meant. The actual mechanics of adjusting for possible confounding were not made clear, and it seemed that different adjustments had been made on different tables.

The possible weaknesses I saw (in addition to some mentioned earlier) were these:

1. Presentation of the data. At times all I had to work with were percents (and I did not know quite what - i.e. what the denominator was) or relative risks. I would have appreciated actual numbers. (For example, in Table 13, it is not clear what the denominators were for analyses 2 and 3).

2. The bimodal age distribution of the controls probably should have led to analyses splitting the ages, rather than of totals. Average age, for example, is not particularly meaningful under the circumstances.

3. The use of first events only makes it difficult to consider continuing or persistent health problems. This, of course, is a recurring problem in morbidity studies - and contributes to the difficulty of comparison with the Canadian Health Survey.

4. It was not clear to me why a proportional mortality approach was necessary (Table 20) for these data. PMR's have some defects and are not as useful as the SMR approach.

* * *

I have serious concern (or suggestions) with only one aspect of tha analysis - that contained in Table 17, "Overlapping Conditions" - the analysis of possible clusters. I believe I understand the rationale for the approach taken by the authors, but I should like to suggest another one. The question asked of cluster-analysis seems to me to be "Do these conditions occur together more often than one would expect?" The key to answering the question is to define what one would expect. I suggest deriving expected completely from the controls - for each event or condition, separately. Thus one would have:

	probability
event i	p_i
event j	p_j

If we assume independence we have:

$$\text{expected } \hat{p}_{i,j} = p_i \times p_j \quad \text{(both)}$$
$$= (1-p_i)\, p_j + p_i\,(1-p_j) \quad \text{(either)}$$
$$= (1-p_i)\,(1-p_j) \quad \text{(neither)}$$

I would now compare observed p_{ij} with expected $\hat{p}_{ij}$, for both exposed and controls, as the ratio $(p_{ij} / \hat{p}_{ij})_E$ and $(p_{ij} / \hat{p}_{ij})_C$

The ratio of these two ratios becomes $\frac{p_{ij}\text{ (exposed)}}{p_{ij}\text{ (controls)}}$

(because the $\hat{p}_{ij}$ is the same for both)

It may be that this ratio will reflect only the more common appearance of conditions i and j in the exposed than in the controls - and hence their more common joint appearance. However, I am concerned that the results now in Table 17 may be an artifact of the computational approach used. (I think I understand why expectations were computed separately for exposed and for controls - and that may be appropriate.)

Here's an extreme example of why I'm worried about the approach in the report:

Event	Probability Exposed	Unexposed
i	.4	.3
j	.3	.2
p_{ij} (expected)	.12	.06

Assert (most extreme case) that every time i appears, if j occurs, j will appear with i (even though j occurs less often). We would then have (observed)

ij	.3	.2
Ratio O/E (within group)	$\frac{.3}{.12} = 2.5$	$\frac{.2}{.06} = 3.33$

$$\text{Ratio } \frac{\text{O/E EXP}}{\text{O/E Control}} = \frac{2.5}{3.33} = .75$$

Thus, if there is a tendency to cluster (here, a strong tendency) the ratio computed in the report should have, or is likely to have a value less than 1 - if more individual events occur among the exposed. This is the artifact that I'm afraid may have occurred.

In fact, if clustering occurs (let us say) 2/3 of the time in the exposed, and 1/2 the time in the controls, this artifact would tend to show no difference - e.g.

Exposed		Controls	
$\frac{.3 \times \frac{2}{3}}{.12}$	$= 1.667$	$\frac{.2 \times \frac{1}{2}}{.06}$	$= 1.667$

The ratio of the two ratios is unity!

* * *

Several details (by page number)

Page:

2 A photograph of the dumpsite would be helpful

4 A housing survey is mentioned (line 3) Is this what is meant? I don't understand

7-8 Comments on Landrigan and Heath are appropriate. I suspect that the L & H criteria are rarely - if ever - met. I see 2 questions for epidemiologic studies to answer:
1. Did an event occur? (to excess)?
2. What led to its occurrence?
I believe Landrigan and Health have mingled the two questions.

9 Long-term-short-term discussions. I find this reasonable. It is more likely that short term effects will also be "subjective".

10 Table 1
I would separate reproductive (relatively easy to observe) from developmental effects (relative hard to observe - at least they take a longer time).

11 line 4 - "metal toxicity". This is not the same sort of thing that is dealt with as neurotoxicity, hepatotoxicity, etc. Need to find another way to say it
line 17 - sister chromatid exchange. It seems to me that this is an appropriate marker even if clinical implications are not known.

12 (typo) - consensus not concensus

13 Two levels of exposure: Higher: regional, salvage picker
Lower: city, firefighter
(I would have liked to have seen more analysis based on this "dose" difference) (More than Table 16)

14 21 deaths - indicate that some analysis is made later on.

14-15 Controls: A problem here - Is it more likely that ill workers would leave the job - and thus not be identified as potential controls?

15 Matching - was matching not possible <u>after</u> the questionnaire was completed - for purposes of analysis?

16 Item 1: These were more <u>common</u> conditions, rather than predominant. How did these reports compare to reports from other landfill studies. Couldn't these conditions have been used to identify items to be looked for - i.e. why weren't these the hypothesis generators? (See item 3, below)

17 Reproductive problems - for males? Or fertility problems for females - which could be dealt with in a study of this kind?

20 Bottom 2 lines - Has a residents' study already been done - as this last sentence seems to indicate? (See also p. 24 middle paragraph)

23 Review by abstractors: See earlier comment about possibilities for quality control, by re-abstracting or comparisons made by another reviewer.

24 (Petty note) There were proportions of confirmations rather than rates. Rates imply some passage of time - or the presence of some time element.

25 Data analysis: Item 1 Indicate that these were adjusted for.
Item 5 (p 27) I don't recall analyses reported with respect to duration of exposure. Were these done and then excluded, or have I misinterpreted Tables 14 and 15?
I think this section would be clearer to the reader if the tables that carried out the proposed analyses were indicated.

28 Lower percent of controls interviewed. Possible serious problem here. Could not "missing" controls be replaced, or new ones recruited?

31 (Table 6) Give distribution of ages by 5 or 10 year age groups. This could provide a framework for later analyses.

32 Tables 8 (p 33) Tables 9 (p 34)
Some of these comparisons should be made by age - as indicated in comment immediately above.
That is, did persons of same ages in exposed and controls have similar distributions of educations attained, or of income? Income is certainly age related.

36 Cigarettes: Table 11

There appears to be rather large differences in age started smoking and the average number of cigarettes smoked. My arithmetic leads me to anticipate 1/4 to 1/3 more lung cancer among the exposed. (Median age started to smoke: 16.5 vs 17.7. Avg. number of cigarettes smoked per respondent 9.4 vs 8.2 These lead me to estimate relative risks by age 50 compared to non-smokers if 6.4 for exposed and 5.1 for controls).

39 Table 12 I think this table needs total cumulative exposures and total number of any exposures (perhaps excluding "metal fumes". What was everyone exposed to that this last % is 98.0 and 97.0?)

Main Analysis - p 40 and following:

Make clear that items 4 (combined respiratory) and 6 (combined skin rash) subsume other items in the table. That is, there are not 13 independent conditions.

Table 13

a) What are the numbers on which Analysis 2 and 3 are based on? 147 and 164 as for analysis 1? Charts were not reviewed on that many persons (see Table 4. p 29)

b) The attrition rate as one goes from Analysis 1 to Analysis 2 and 3, with only two exceptions, items 7 (angina, heart attack), and 10 (anxiety and depression) greater for controls than for exposed. What explanation? It seemed to me that a later table (Table 25) showed, in general, greater confirmation proportions for controls. Table 26 seems also to relate to this aspect of Table 13, but I am unable to fit them together.

See remarks on p 44 item 6 about loss of exposure-illness association for angina, heart attack. This was one of two items in which attrition for exposed was greater than for controls.

44 Paragraph beginning "No association. " These data might be better displayed in a table.

45 First paragraph, last sentence. This needs to be explained.

46 Table 14 The "condition " descriptions are different than those in Table 13. They should be the same.

47 and 48 Table 15 Age adjusted? Same comment re: descriptions.

49 Item 1 This is a true defect in this study - and is likely to be a problem in the residents' study as well.

Item 3 last paragraph - common factor. I see no way to avoid it.

50, 51 Table 16
Could the controls have been used as still another level of exposure - giving 3 "doses" rather than 2?
One could ask what is the probability that 11 of 13 (or 9 of 11, if one leaves out overlaps) show a "dose" gradient if hypothesis of no effect implies half should show increases with dose, and half should show decrease with increasing dose.

52, 53 Table 17 - discussed earlier

55 56 Table 18 - I would also show the odd ratios (Analysis 1?) from Table 13, and the p values. Then it becomes necessary to explain why p values declined (Table 18 < Table 13) for 2 items, #7, and #12, and increased for all others. Do some measures of acute effect (red, itchy eyes, skin rash?) provide an estimate of over-reporting by exposed (previously exposed) persons- or only a measure of persistence of the effect? Condition descriptions again.

57 Table 19 Any explanation needed for higher percent of Workmen's Compensation awards among exposed? Is double spontaneous abortions among exposed meaningful - or were

numbers too small. What was relative risk, or odds ratio?

Mortality study - mentioned earlier.

Canada Health Survey - mentioned earlier

Some data gathering problems must be present:
Controls: all odds ratios but 2 exceed unity!

63 Table 22 - COntrols were "outdoor" workers. Do persons with hayfever and/or asthma seek outdoor employment (away from dusts, fumes, etc. of a landfill?) Does this help explain odds ratios of less than unity? 12 of 18 items have P's of less than .2. If there were no associations we would have expected 3 or 4.

65 Table 23 - Same data by exposed and controls?

66 Table 24 - Note "confirmed or possible confirmed"

pre 1965	99%!
1965-69	89.7
1970-75	88.9
1976-80	85.8
1981-83	85.3

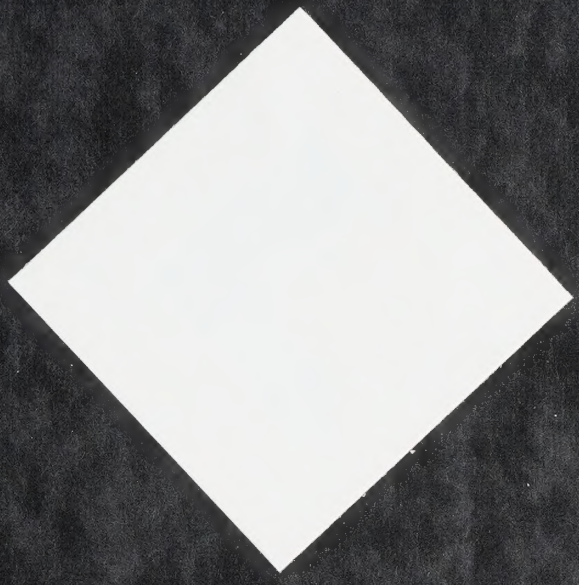
62-68
70-71 Tables 25, 26 Commented on above

72 Table 27 - physicians attendance. Need to explain why 11/12 show controls with better (or more) use of physician. Is this an indication of over-reporting by exposed?

74 Table 28 This table is really the essence of the report. Some development (references, etc) is needed for classification under biological plausibility. A supplementary tabulation may be in order for effects believed to be biologically plausible (or epidemiologically identified or suggested) that did not show association with exposure. (Perhaps "bleeding", urinary tract infections - and other things reported in the intitial residents' survey)

Additional general comment (or, perhaps errant thought)

If biological plausibility can be set out in advance, should one then use some Bayesian approach to evaluating results? If I have a prior probability that an event should (or will) occur - and it does occur should that not lend strength to the postulated association?



ACCOPRESS®

25071 –	BLACK / NOIR	– BG2507
25072 –	BLUE / BLEU	– BU2507
25078 –	RED / ROUGE	– BF2507
25075 –	GREEN / VERT	– BP2507
25074 –	GREY / GRIS	– BD2507
25073 –	R. BLUE / BLEU R.	– BB2507
25079 –	X. RED / ROUGE X.	– BX2507
25070 –	YELLOW / JAUNE	– BY2507
25077 –	TANGERINE	– BA2507

ACCO CANADIAN COMPANY LIMITED
COMPAGNIE CANADIENNE ACCO LIMITÉE
TORONTO CANADA

HAMILTON PUBLIC LIBRARY
3 2022 21294054 4